HZCYTOR02.SEQ - ▮▮▮▮▮▮▮

Sequence of pcr products generated with 9800-9802,
nested pcr product 9941-AP2 (9801-AP1)
nested pcr product 9937-AP2 (9803-AP1)

Enzyme	Recognition	Cut Site
AgeI	(A^CCGGT)	Def: 1124
BamHI	(G^GATCC)	Def: 172
DraI	(TTT^AAA)	Def: 36
EcoRI	(G^AATTC)	Def: 450
EcoRV	(GAT^ATC)	Def: 438
HpaI	(GTT^AAC)	Def: 145
MscI	(TGG^CCA)	Def: 1244
MunI	(C^AATTG)	Def: 493
NcoI	(C^CATGG)	Def: 377
NsiI	(ATGCA^T)	Def: 592
Ppu10I	(A^TGCAT)	Def: 588
SmaI	(CCC^GGG)	Def: 11
SspI	(AAT^ATT)	Def: 503 988 1107
XmaI	(C^CCGGG)	Def: 9

This should be the full sequence. The other sequence I gave you had 3 mistakes in it.

HZCYTOR02.SEQ   Linear   LENGTH = 1289

```
          XmaI
          SmaI                    DraI            4q
          | |                     |               |
  1 CCCCCCGCCCGGGAGAGAGGCAATATCAAGGTTTTAAATCTCGGAGAAATGGCTTTCGTTTGCTTGGCT   69
    GGGGGGCGGGCCCTCTCTCCGTTATAGTTCCAAAATTTAGAGCCTCTTTACCGAAAGCAAACGAACCGA   7
          |                       |               M  A  F  V  C  L  A
          11                      36
          9
```

```
 70 ATCGGATGCTTATATACCTTTCTGATAAGCACAACATTTGGCTGTACTTCATCTTCAGACACCGAGATA   138
    TAGCCTACGAATATATGGAAAGACTATTCGTGTTGTAAACCGACATGAAGTAGAAGTCTGTGGCTCTAT   30
     I  G  C  L  Y  T  F  L  I  S  T  T  F  G  C  T  S  S  S  D  T  E  I
                                                       Prd. N-term S25
```

```
        HpaI                    BamHI
        |                       |
139 AAAGTTAAACCCTCCTCAGGATTTTGAGATAGTGGATCCCGGATACTTAGGTTATCTCTATTTGCAATGG   207
    TTTCAATTGGGAGGAGTCCTAAAACTCTATCACCTAGGGCCTATGAATCCAATAGAGATAAACGTTACC   53
     K  V  N  P  P  Q  D  F  E  I  V  D  P  G  Y  L  G  Y  L  Y  L  Q  W
        145                     172
```

```
208 CAACCCCCCACTGTCTCTGGATCATTTTAAGGAATGCACAGTGGAATATGAACTAAAATACCGAAACATT   276
    GTTGGGGGTGACAGAGACCTAGTAAAATTCCTTACGTGTCACCTTATACTTGATTTTATGGCTTTGTAA   76
     Q  P  P  L  S  L  D  H  F  K  E (C) T  V  E  Y  E  L  K  Y  R  N  I
```

```
277 GGTAGTGAAACATGGAAGACCATCATTACTAAGAATCTACATTACAAAGATGGGTTTGATCTTAACAAG   345
    CCATCACTTTGTACCTTCTGGTAGTAATGATTCTTAGATGTAATGTTTCTACCCAAACTAGAATTGTTC   99
     G  S  E  T  W  K  T  I  I  T  K  N  L  H  Y  K  D  G  F  D  L  N  K
```

```
                        NcoI
                        |
346 GGCATTGAAGCGAAGATACACACGCTTTTACCATGGCAATGCACAAATGGATCAGAAGTTCAAAGTTCC   414
    CCGTAACTTCGCTTCTATGTGTGCGAAAATGGTACCGTTACGTGTTTACCTAGTCTTCAAGTTTCAAGG   122
     G  I  E  A  K  I  H  T  L  L  P  W  Q  C  T  N  G  S  E  V  Q  S  S
                        377
```

```
                EcoRV          EcoRI
                |              |
415 TGGGCAGAAACTACTTATTGGATATCACCACAAGGAATTCCAGAAACTAAAGTTCAGGATATGGATTGC   483
    ACCCGTCTTTGATGAATAACCTATAGTGGTGTTCCTTAAGGTCTTTGATTTCAAGTCCTATACCTAACG   145
     W  A  E  T  T  Y  W  I  S  P  Q  G  I  P  E  T  K  V  Q  D  M  D  C
                438            450
```

```
        MunI      SspI
        |         |
484 GTATATTACAATTGGCAATATTTACTCTGTTCTTGGAAACCTGGCATAGGTGTACTTCTTGATACCAAT   552
    CATATAATGTTAACCGTTATAAATGAGACAAGAACCTTTGGACCGTATCCACATGAAGAACTATGGTTA   168
     V  Y  Y  N  W  Q  Y  L  L  C  S  W  K  P  G  I  G  V  L  L  D  T  N
        493       503
```

23 a-al line

Mature 25-380

Tmd = 25
EK-cell - 25 - 339
Tmd 340-363
Cyto 364 -380

```
                                     Ppu10I
                                       NsiI
                                       |   |
553 TACAACTTGTTTTTACTGGTATGAGGGCTTGGATCATGCATTACAGTGTGTTGATTACATCAAGGCTGAT 621
    ATGTTGAACAAAATGACCATACTCCCGAACCTAGTACGTAATGTCACACAACTAATGTAGTTCCGACTA   191
    Y  N  L  F  Y  W  Y  E  G  L  D (H) A  L  Q  C  V  D  Y  I  K  A  D
                                       588    592


622 GGACAAAATATAGGATGCAGATTTCCCTATTTGGAGGCATCAGACTATAAAGATTTCTATATTTGTGTT 690
    CCTGTTTTATATCCTACGTCTAAAGGGATAAACCTCCGTAGTCTGATATTTCTAAAGATATAAACACAA   214
    G  Q  N  I  G  C  R  F  P  Y  L  E  A  S  D  Y  K  D  F  Y  I  C  V


691 AATGGATCATCAGAGAACAAGCCTATCAGATCCAGTTATTTCACTTTTCAGCTTCAAAATATAGTTAAA 759
    TTACCTAGTAGTCTCTTGTTCGGATAGTCTAGGTCAATAAAGTGAAAAGTCGAAGTTTTATATCAATTT   237
    N  G  S  S  E  N  K  P  I  R  S  S  Y  F  T  E  Q  L  Q  N  I  V  K


760 CCTTTGCCGCCAGTCTATCTTACTTTTTACTCGGGAGAGTTCATGTGAAATTAAGCTGAAATGGAGCATA 828
    GGAAACGGCGGTCAGATAGAATGAAAATGAGCCCTCTCAAGTACACTTTAATTCGACTTTACCTCGTAT   260
    P  L  P  P  V  Y  L  T  F  T  R  E  S  S  C  E  I  K  L  K  W (S) I


829 CCTTTGGGACCTATTCCAGCAAGGTGTTTTGATTATGAAATTGAGATCAGAGAAGATGATACTACCTTG 897
    GGAAACCCTGGATAAGGTCGTTCCACAAAACTAATACTTTAACTCTAGTCTCTTCTACTATGATGGAAC   283
    P  L  G  P  I  P  A  R  C  F  D  Y  E  I  E  I  R  E  D  D  T  T  L


898 GTGACTGCTACAGTTGAAAATGAAACATACACCTTGAAAACAACAAATGAAACCCGACAATTATGCTTT 966
    CACTGACGATGTCAACTTTTACTTTGTATGTGGAACTTTTGTTGTTTACTTTGGGCTGTTAATACGAAA   306
    V  T  A  T  V  E  N  E  T  Y  T  L  K  T  T  N  E  T  R  Q  L  C  F


                   SspI
                    |
967 GTAGTAAGAAGCAAAGTGAATATTTATTGCTCAGATGACGGAATTTGGAGTGAGTGGAGTGATAAACAA 1035
    CATCATTCTTCGTTTCACTTATAAATAACGAGTCTACTGCCTTAAACCTCACTCACCTCACTATTTGTT   329
    V  V  R  S  K  V  N  I  Y  C  S  D  D  G  I  W  S  E  W  S  D  K  Q
                    988


1036 TGCTGGGAAGGTGAAGACCTATCGAAGAAAACTTTGCTACGTTTCTGGCTACCATTTGGTTTCATCTTA 1104
     ACGACCCTTCCACTTCTGGATAGCTTCTTTTGAAACGATGCAAAGACCGATGGTAAACCAAAGTAGAAT   352
     C  W  E  G  E  D  L  S  K  K  T  L  L  R  F  W  L  P  F  G  F  I  L
                          337         TMD


     SspI           AgeI
      |              |
1105 ATATTAGTTATATTTGTAACCGGTCTGCTTTTGCGTAAGCCAAACACCTACCCAAAAATGATTCCAGAA 1173
     TATAATCAATATAAACATTGGCCAGACGAAAACGCATTCGGTTTGTGGATGGGTTTTTACTAAGGTCTT   375
     I  L  V  I  F  V  T  G  L  L  L  R  K  P  N  T  Y  P  K  M  I  P  E
     1107            1124  363


1174 TTTTTCTGTGATACATGAAGACTTTCCATATCAAGAGACATGGTATTGACTCAACAGTTTCCAGTCATG 1242
     AAAAAGACACTATGTACTTCTGAAAGGTATAGTTCTCTGTACCATAACTGAGTTGTCAAAGGTCAGTAC
     F  F  C  D  T  .


     MscI
      |
1243 GCCAAATGTTCAATATGAGTCTCAATAAACTGAATTTTTCTTGCGAA 1289
     CGGTTTACAAGTTATACTCAGAGTTATTTGACTTAAAAAGAACGCTT

     1244
```